

## Aberystwyth University

### *Statistical Modeling for Spatial Groundwater Potential Map Based on GIS Technique*

Azma, Aliasghar; Narreie, Esmaeil; Shojaaddini, Abouzar; Kianfar, Nima; Kiyanfar, Ramin; Alizadeh, Seyed Mehdi Seyed; Davarpanah, Afshin

*Published in:*  
Sustainability (Switzerland)

*DOI:*  
[10.3390/su13073788](https://doi.org/10.3390/su13073788)

*Publication date:*  
2021

*Citation for published version (APA):*

Azma, A., Narreie, E., Shojaaddini, A., Kianfar, N., Kiyanfar, R., Alizadeh, S. M. S., & Davarpanah, A. (2021). Statistical Modeling for Spatial Groundwater Potential Map Based on GIS Technique. *Sustainability (Switzerland)*, 13(7), [3788]. <https://doi.org/10.3390/su13073788>

#### **Document License** CC BY

#### **General rights**

Copyright and moral rights for the publications made accessible in the Aberystwyth Research Portal (the Institutional Repository) are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Aberystwyth Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Aberystwyth Research Portal

#### **Take down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

tel: +44 1970 62 2400  
email: [is@aber.ac.uk](mailto:is@aber.ac.uk)

Article

# Statistical Modeling for Spatial Groundwater Potential Map Based on GIS Technique

Aliasghar Azma <sup>1</sup> , Esmail Narreie <sup>2</sup>, Abouzar Shojaaddini <sup>3</sup>, Nima Kianfar <sup>4,\*</sup>, Ramin Kiyanfar <sup>5</sup> ,  
Seyed Mehdi Seyed Alizadeh <sup>6</sup> and Afshin Davarpanah <sup>7</sup> 

- <sup>1</sup> College of Architecture and Civil Engineering, Beijing University of Technology, Beijing 100124, China; aa.azma1990@emails.bjut.edu.cn
- <sup>2</sup> Department of Surveying Engineering, Faculty of Civil and Surveying Engineering, Graduate University of Advanced Technology, Kerman 76311-33131, Iran; e.narreie@student.kgut.ac.ir
- <sup>3</sup> Soil Science Department, College of Agriculture, Tarbiat Modares University, Tehran 14115-336, Iran; a.shojaaddini68@yahoo.com
- <sup>4</sup> Faculty of Geodesy & Geomatics Engineering, K. N. Toosi University of Technology, Tehran 15875-4416, Iran
- <sup>5</sup> Department of Art and architecture, Payame Noor University, Shiraz 19395-4697, Iran; raminkianfar@spnu.ac.ir
- <sup>6</sup> Petroleum Engineering Department, Australian College of Kuwait, West Mishref 13015, Kuwait; s.alizadeh@ack.edu.kw
- <sup>7</sup> Department of Mathematics, Aberystwyth University, Aberystwyth SY23 3FL, UK; afd6@aber.ac.uk
- \* Correspondence: Nimakianfar@email.kntu.ac.ir



**Citation:** Azma, A.; Narreie, E.; Shojaaddini, A.; Kianfar, N.; Kiyanfar, R.; Seyed Alizadeh, S.M.; Davarpanah, A. Statistical Modeling for Spatial Groundwater Potential Map Based on GIS Technique. *Sustainability* **2021**, *13*, 3788. <https://doi.org/10.3390/su13073788>

Academic Editors: Ozgur Kisi and Ombretta Paladino

Received: 25 January 2021  
Accepted: 19 March 2021  
Published: 29 March 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Abstract:** In arid and semi-arid lands like Iran water is scarce, and not all the wastewater can be treated. Hence, groundwater remains the primary and the principal source of water supply for human consumption. Therefore, this study attempted to spatially assess the groundwater potential in an aquifer in a semi-arid region of Iran using geographic information systems (GIS)-based statistical modeling. To this end, 75 agricultural wells across the Marvdasht Plain were sampled, and the water samples' electrical conductivity (EC) was measured. To model the groundwater quality, multiple linear regression (MLR) and principal component regression (PCR) coupled with eleven environmental parameters (soil-topographical parameters) were employed. The results showed that that soil EC (SEC) with Beta = 0.78 was selected as the most influential factor affecting groundwater EC (GEC). CaCO<sub>3</sub> of soil samples and length-steepness (LS factor) were the second and third effective parameters. SEC with  $r = 0.89$  and CaCO<sub>3</sub> with  $r = 0.79$  and LS factor with  $r = 0.69$  were also characterized for PC1. According to performance criteria, the MLR model with  $R^2 = 0.94$ , root mean square error (RMSE) = 450  $\mu\text{Scm}^{-1}$  and mean error (ME) = 125  $\mu\text{Scm}^{-1}$  provided better results in predicting the GEC. The GEC map indicated that 16% of the Marvdasht groundwater was not suitable for agriculture. It was concluded that GIS, combined with statistical methods, could predict groundwater quality in the semi-arid regions.

**Keywords:** carbonate aquifer; digital elevation model; modeling; multivariate linear regression; principal component regression; groundwater quality assessment

## 1. Introduction

Groundwater is the foremost valuable water resource worldwide, particularly in arid and semi-arid regions like Iran. In addition to the lack of surface water resources, due to the lack of wastewater treatment technology in many places in Iran, it is impossible to use the wastewater. [1]. In provinces and territories located in the south of Iran with an arid and semi-arid climate, groundwater is the primary and sometimes the only resource for fresh water [2–4]. The demand for groundwater supplies for human consumption, including drinking, agriculture, industry, etc., has widely increased with the increasing human population in recent decades, with little opportunity to treat the wastewater. Hence, groundwater quality analysis is vital to manage the groundwater reservoirs [1,3,5].

Therefore, in order to obtain an exact measurement of groundwater quality, conventional methods should be generally used for various groundwater-related parameters, along with their complexity. These methods require lots of time, money, and labor [1,5]. Furthermore, groundwater sampling is difficult and not even possible in particular regions. Additionally, groundwater quality generally is connected to topography, geological structures, lithology, soil, and slope [6]. Therefore, newly developed big data set methods need to be employed to analyze the groundwater, as it is simply assessed using statistical models for the interpretation of the environmental data [1,6].

In most areas of Iran, especially in the southern regions, due to the presence of carbonate formations, groundwater pH changes very little, and therefore it cannot be an accurate indicator of the potential of groundwater quality in the regions; however, it has been used as an indicator in many types of research. Hence, groundwater salinity (electrical conductivity; EC) can be a better indicator to show the potential of groundwater quality, particularly in areas with high EC variation [1,5]. As the only analysis of groundwater quality data is not sufficient to assess the groundwater quality for different purposes because of such a big data set, particularly at watershed scale with numerous sampling wells, statistical models can cope with these issues properly. Principal component regression (PCR) and multiple linear regression (MLR) are two applicable methods and have been applied to investigate the temporal-spatial analysis of groundwater quality potential [6–9]. The PCR allows identifying significant quality parameters, and MLR is used to find the relationship between the parameters and develop the prediction models. MLR can also determine the relative influence of one or more predictor variables and identify outliers or anomalies [6,10]. Geographic information systems (GIS) as a visual tool provides applicable methods for manipulating spatial data [11] because these tools have prepared a time- and cost-saving way for groundwater quality evaluation [12,13]. Hence, groundwater quality mapping integrated with statistical modeling has been widely used for carrying out successful groundwater protection and management programs [1–3]. Many studies have used the combination of GIS and statistical methods for the spatial groundwater quality analysis [3,14,15]. Haghizadeh et al. [2] applied GIS integrated with statistical techniques for analyzing the Broujerd region's groundwater in Iran. They used 11 factors extracted from the digital elevation model (DEM) for potential groundwater mapping. Yadav et al. [16] used a GIS-based approach combined with principal component analysis (PCA) analysis to evaluate groundwater geochemistry and statistical determination of the fate of contaminants in shallow aquifers from different functional areas of Agra city, India. Their results showed that the combination of PCA and GIS is beneficial and provides acceptable results for assessing groundwater contamination. Naghibi et al. [17] used different types of data-mining methods and 13 hydrological-geological-physiographical factors for producing the potential groundwater maps in Koohrang, Iran. Their results showed that environmental factors integrated with the data-mining methods could help map the potential groundwater quality. Arulbalaji et al. [15] used GIS- analytic hierarchy process (AHP) to assess the groundwater quality potential in southern Western Ghats, India. They used 12 thematic layers such as geology, geomorphology, soil, and slope for groundwater potential zone demarcation. Their results showed that the accuracy of the method was around 85%. Mosafari et al. [18] used multivariate statistical techniques, including principal component analysis (PCA) and hierarchical cluster analysis (CA) coupled with GIS, to investigate the groundwater quality for drinking purposes in rural northwest Iran. They showed that the multivariate analysis within GIS could be applicable for the assessment of groundwater in Iran. Honarbakhsh et al. reported that the geostatistical analysis combined with GIS provides acceptable groundwater quality assessment results by investigating the Marvdasht Aquifer [3]. Abdalla et al. [19] revealed that integrating remote sensing data and GIS methods gives a valuable tool for the improved prediction, monitoring, and planning of water resources in arid and hyper-arid regions. Ijumulana et al. [20] used GIS for mapping the hotspots and potential health risks of groundwater in Pakistan. They reported that the probability of having a safe source of drinking water varied from one

geological unit to another, with sources in the Neogene Quaternary volcanic formations having the least probabilities.

Although GIS-based mapping of groundwater quality is beneficial for preserving and managing this resource, only a few works used GIS techniques integrated with statistical models to evaluate Iran's groundwater quality potential. However, there have been no published works in Iran, particularly in southern Iran, to investigate the relationship between soil-topographical parameters and groundwater quality using statistical modeling approaches coupled with GIS. In addition, the effect of geological formation on groundwater quality has not been deeply investigated. Hence, the current work aimed to: (1) investigate the application of two statistical modeling approaches (PCR and MLR) for the assessment of the groundwater quality map of Marvdasht Aquifer as the most important aquifer in the Fars because it supplies drinking water for more than 300,000 people and irrigation water for agricultural activities.; (2) test the use of soil-related parameters and topography attributes as model inputs for evaluating the potential of groundwater quality assessment combined with statistical methods; and (3) investigate the performance of GIS application with statistical models for the spatial groundwater potential analysis.

## 2. Materials and Methods

### 2.1. Study Site and Sampling

The site (3926.3 km<sup>2</sup>) is situated between 29°19'–30°20' N and 52°15'–53°27' E in the Marvdasht region, Fars, Iran (Figure 1). The aquifer covers an area of 1986.4 km<sup>2</sup> (Figure 1). The Marvdasht Aquifer supplies drinking water for Marvdasht city and around 200 villages with more than 300,000 inhabitants. It also supplies fresh water for people and agricultural purposes [3,4]. The site has a semi-arid climate with a mean rainfall and temperature of 291.7 mm and 17.5 °C, respectively.

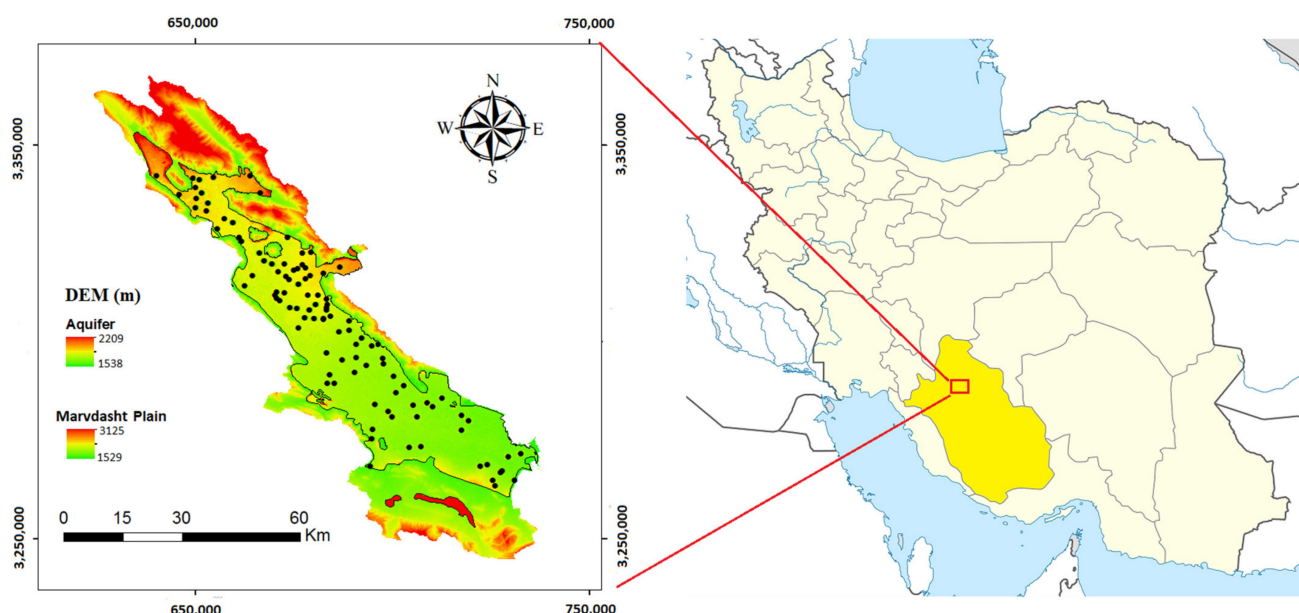
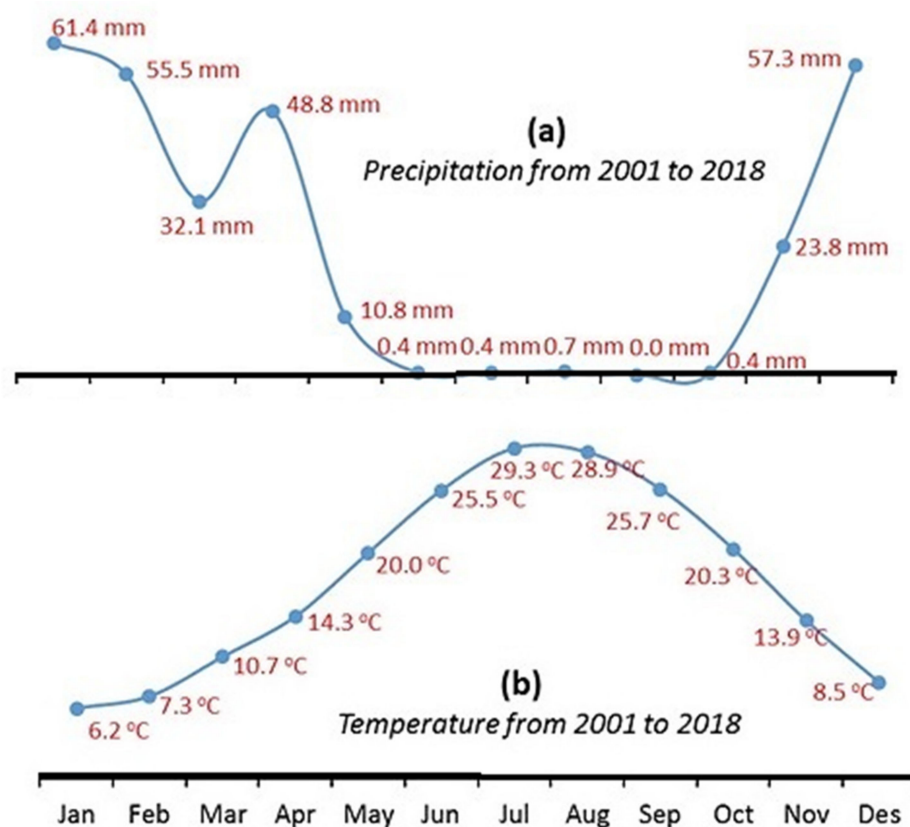


Figure 1. Location of the study site and sampling wells.

The most rainfall happens during December, January, and February (Figure 2a). Temperatures are high during June, July, August, and September and relatively low during the winter months (Figure 2b). The altitude ranges from 1529 to 3125 m above the sea. The center of the site is flat under agricultural farms, while the elevated parts are mainly mountainous. The prominent geological formations are dolomite and calcite of Sarvak, conglomerates, and alluvial deposits related to Q1 to Q3 [4,17]. Moreover, Hormuz and Asmari have soluble materials, including chalky–salty marl and limestone [2,21]. Soils

are calcareous, including Inceptisols, Entisols, and Aridisols (Soil Taxonomy) [22,23]. The water table is around 35 m [1,24].



**Figure 2.** Data for precipitation (a) and temperature (b) in the study site.

As the Marvdasht Plain is the most important agricultural area in the south of Iran and groundwater provides more than 85% of the irrigation water in this region, we used electrical conductivity (EC) of 75 agricultural wells (Figure 1) to model the groundwater quality potential. The EC was measured using the portable EC meter. Locations of the wells were recorded using a GPS. Additionally, close to the studied wells, approximately 1 kg topsoil was taken (90 soil samples in total). The air-dried, ground, and sieved soil samples were used for lab analysis, including soil texture by the hydrometer method [20], calcium carbonate by neutralization with 1 N HCl [25], and electrical conductivity (EC) using the EC meter.

## 2.2. Methodology

The flow chart of the methodology is presented in Figure 3. The map of soil parameters was depicted using the ordinary kriging (OK), which has been widely used for groundwater quality assessment and showed good performance in assessing the groundwater quality parameters [16], and variogram in ArcGIS 10.6 as follows [26].

$$Y(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(x_i) - Z(x_i + h)]^2, \quad (1)$$

$$\hat{Z}(x_i) = \sum_{i=1}^N \lambda_i Z(x_i) \quad (2)$$

where  $Z(x_i)$ ,  $Y(h)$ , and  $N(h)$  are variables for a location  $x_i$ , the variogram for a lag distance  $h$  between  $Z(x_i)$  and  $Z(x_i + h)$  and data pairs, respectively. The following topographical



factors were identified as essential input maps/layers (predictor variables) for the GIS-based model development of groundwater quality [26,27].

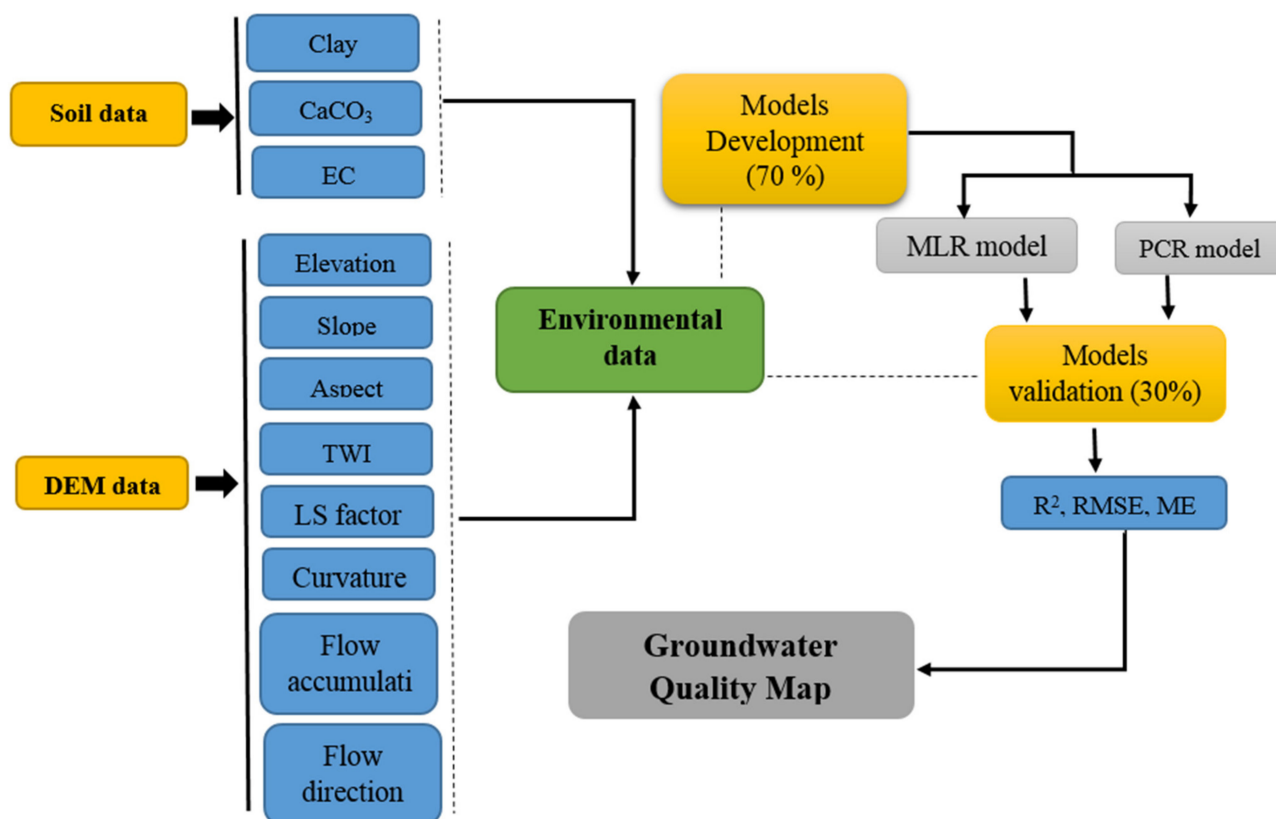


Figure 3. Flow chart of the methodology used.

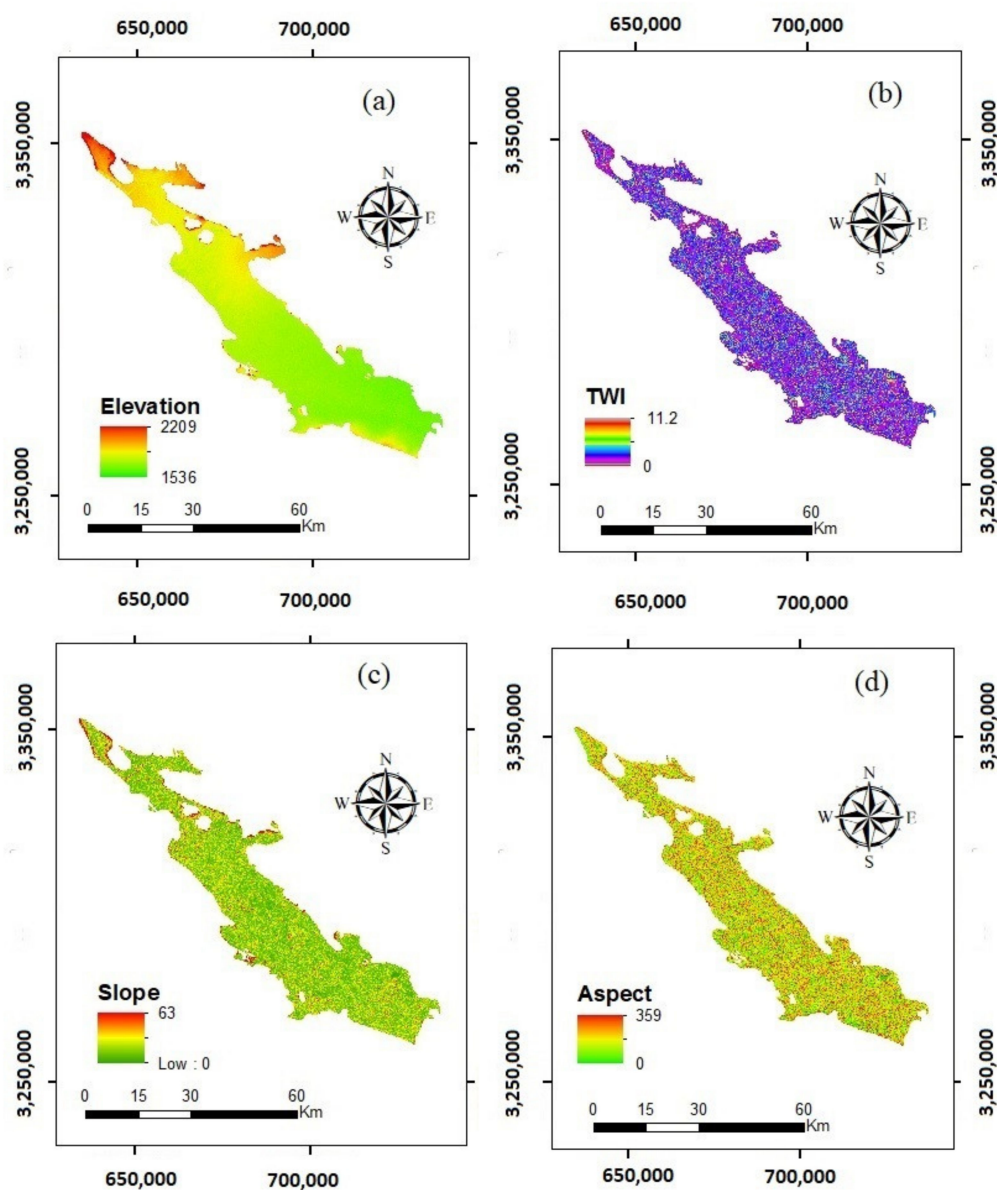
Digital Elevation Model (DEM) with a 30-m cell size was used for extracting the layers of the slope, altitude, aspect, topographic wetness index (TWI), slope length and steepness (LS factor), flow accumulation, flow direction, and curvature, which are depicted in Figure 4. These contributing layers have been frequently used for groundwater quality assessment worldwide and particularly in Iran. The elevation (altitude) of the Marvdasht Aquifer varied from 1536 to 2209 m above sea level (Figure 4a). The TWI was obtained from DEM in SAGA GIS with a range of 0 to 11.2 (Figure 4b). TWI is used to describe the impact of topography on runoff [28], which is calculated as follows (Moore and Burch 1986):

$$TWI = \ln \left( \frac{FA}{\tan \beta} \right). \quad (3)$$

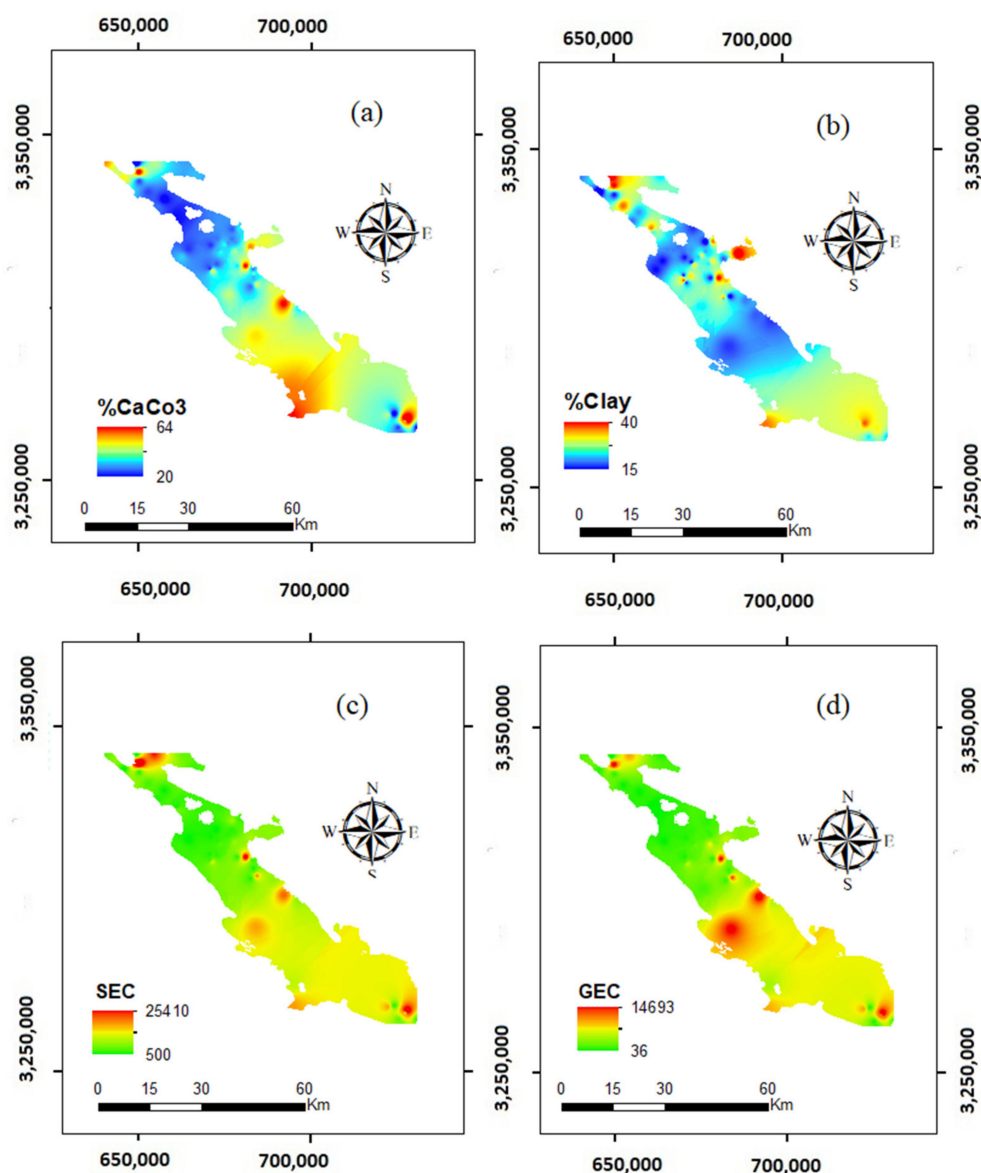
TWI is the topographic wetness index; FA is flow accumulation that describes the accumulated up-slope-related area for a given cell. Flow accumulation was created from the flow direction raster map created from DEM that determined the direction of the water flow in a given cell. The aquifer's slope was created from DEM in ArcMap 10.6 and ranged from 0 to 63 degrees (Figure 4c). The LS factor describes the slope length and steepness. The LS factor is considered for quantifying soil erosion [28]. The LS factor map was calculated using the equation suggested by Moore and Burch [26] and the ArcGIS 10.6 spatial analyst extension with a DEM as follows:

$$LS = \left( \frac{FA \times \text{Cell size}}{22.13} \right)^{0.4} \times \left( \frac{\sin \text{slope}}{0.0896} \right)^{1.3}. \quad (4)$$

LS is the slope length and slope steepness combined, as outlined by Moore and Burch (1986), FA is flow accumulation, and sin slope is the slope map (Figure 4c), and aspect map (Figure 4d) derived from the DEM. The curvature map was created from DEM in SAGA GIS software. Figure 5 shows the maps of soil clay, soil carbonate calcium ( $\text{CaCO}_3$ ), soil EC (SEC), and groundwater EC (GEC). As shown in Figure 5a, soil  $\text{CaCO}_3$  varied from 20% in the north of the study site (Marvdasht Aquifer) to 64% south of the study site (Figure 5a). Soil clay varied in the range of 15–40% (Figure 5b). There was a very similar SEC (Figure 5c) and GEC (Figure 5d) spatial distribution in the study site. The EC's highest soil and groundwater values were found in an area with soluble materials [1,3]. The values of topography parameters (Figure 4a–d) at 75 studied wells extracted in ArcMap 10.6 and soil data (clay,  $\text{CaCO}_3$ , and EC) from those portions were used as input variables in model development.



**Figure 4.** Topographical maps of the site: (a) elevation, (b) topographic wetness index (TWI), (c) slope and (d) aspect.



**Figure 5.** Maps of soil carbonate calcium (a), soil clay (b), soil electrical conductivity (SEC) (c), and groundwater EC (GEC) (d).

Before modeling, all topographic variables were standardized to deal with the problem of non-uniform units. Additionally, before modeling, the normality test of the data was done. Seventy-five wells were randomly divided into 53 wells (70%) for model development and 22 wells (30%) for model validation. The *t*-test was used to check the differences between two data sets (validation and calibration data set) with a significance level of 0.05. Pearson correlation analyses with a significance level of 0.05 between the EC and topographic variables and soil parameters were implemented. The significant topographical variables were then selected to be incorporated into the stepwise multiple linear regressions (SMLR) model developed to predict the groundwater EC. Then, the forward stepwise multiple linear regressions (MLRs) were conducted to advance the model in order to predict GEC as follows:

$$\text{GEC} = A + a_1 X_1 + a_2 X_2 + a_3 X_3 + \dots + a_N X_N. \quad (5)$$

The GEC is the groundwater's electrical conductivity; *A* is the intercept,  $X_1$  to  $X_N$  are input/independent variables (soil-topographical parameters), and  $a_1$  to  $a_8$  are regression



coefficients. Multicollinearity [24] was checked by the Variance Inflation Factor (VIF) as follows [15]:

$$\text{VIF} = \frac{1}{1 - R_j^2} \quad (6)$$

where  $R_j^2$  is adjusted  $R^2$ , a  $\text{VIF} > 5$  indicates collinearity among input variables [29]. The MLR was fulfilled in the Statistica 8.0 software.

Principal component regression (PCR) was carried out using the PCR module in the Statistica 8.0 software to avoid multicollinearity, reduce the dimensions of the data set, and decrease the input number of variables [30]. The principal components (PCs) that accounted for at least 90% of the original data set variance were selected [30,31]. Then, the PCs were applied as input variables (independent variables). Equation (5) was employed to rescale the EC and input variables.

$$X_n = 0.1 + 0.8 \left[ \frac{X - X_{\min}}{X_{\max} - X_{\min}} \right] \quad X_{\min} < X < X_{\max}. \quad (7)$$

$X_n$  is the rescaled data;  $X_{\min}$  and  $X_{\max}$  are the min and max values of observed data.

The MLR and PCR were evaluated using  $R^2$ , the mean error (ME), and the root mean square error (RMSE) as follows:

$$R^2 = \frac{[\sum_{i=1}^n (O_i - \bar{O}) (P_i - \bar{P})]^2}{\sum_{i=1}^n (O_i - \bar{O})^2 \sum_{i=1}^n (P_i - \bar{P})^2}, \quad (8)$$

$$\text{RMSE} = \sqrt{\left( \frac{\sum_{i=1}^n (P_i - O_i)^2}{n} \right)}, \quad (9)$$

$$\text{ME} = \frac{\sum_{i=1}^n (P_i - O_i)}{n} \quad (10)$$

where  $n$  is the number of observations,  $O$  is the observed values,  $P$  is the predicted values, and  $i$  is the number of samples [31].

### 3. Results

#### 3.1. Soil Descriptive and Geostatistical Analysis

A summary of soil properties (texture,  $\text{CaCO}_3$ , and EC) is given in Table 1. A negative skewness was observed for EC and the silt. In contrast, a little positive skewness was identified for others. However, the Kolmogorov–Smirnov test showed normally distributed properties. In addition, a coefficient of variation (CV) test proved that properties with a  $\text{CV} > 35\%$  had high variability, those with a  $\text{CV} < 35\%$  had moderate variability, and those with a  $\text{CV} < 15\%$  had low variability.

**Table 1.** Descriptive statistics of study soils.

Properties	Min	Max	Mean	Median	Std. dev	CV
Clay (%)	15.1	40.0	26.5	27.3	4.53	17.1
Silt (%)	24.0	40.0	31.7	31.6	3.20	10.1
Sand (%)	30.1	55.0	41.8	42.1	4.94	11.8
$\text{CaCO}_3$ (%)	20.0	64.0	43.5	45.0	12.05	27.7
EC ( $\mu\text{Scm}^{-1}$ )	500	25,410	2465	2350	224	93.3

CV: Coefficient of variation;  $\text{CaCO}_3$ : Calcium carbonate; EC: Electrical conductivity.

The videography analysis of soil properties is presented in Tables 2 and 3. All variograms for all soil properties except silt were anisotropic. The spherical and Gaussian models were the best for soil properties (Table 2).

**Table 2.** Semi-variograms for soil and groundwater EC.

Parameters	Model	Range (km)		k	Direction (Degree)	Nugget	Sill	Nugget/Sill
		Minor	Major					
SEC	Gaussian	19.19	23.89	1.24	56.6	2.29	3.79	0.604
CaCO <sub>3</sub>	Spherical	21.46	43.49	2.02	97.7	2.08	4.09	0.508
GEC	Spherical	17.54	32.59	1.86	118.1	3.84	11.12	0.345
Clay	Gaussian	32.56	51.00	1.57	66.2	8.89	18.43	0.482
Sand	Gaussian	25.70	45.37	1.76	133.1	6.81	23.31	0.292

**Table 3.** Ordinary kriging information of soil properties and groundwater EC.

Parameters	Interpolation Method	R <sup>2</sup>	RMSE	ME
SEC	OK	0.79	2.145	0.014
CaCO <sub>3</sub>	OK	0.65	2.318	0.044
GEC	OK	0.55	0.991	0.143
Clay	OK	0.78	4.244	−0.024
Sand	OK	0.72	4.371	−0.071

### 3.2. MLR Model Development and Validation

The relationships between groundwater EC (GEC) and soil-topographical parameters are given in Table 4.

**Table 4.** Correlation of soil-topographical parameters with EC of the groundwater (\* represents significant level,  $p < 0.05$ ).

	Elevation	Slope	Aspect	TWI	Flow Di- rection	LS Factor	Flow Length	SEC	CaCO <sub>3</sub>	Clay
Slope	0.18									
Aspect	−0.23	−0.13								
TWI	0.10	0.26 *	0.11							
Flow Direction	−0.20	−0.19	−0.16	0.04						
LS factor	0.07	0.14	−0.31 *	−0.44 *	−0.11					
Flow length	−0.06	−0.03	−0.18	−0.13	−0.08	0.12				
SEC	−0.22	0.18	0.04	−0.12	−0.07	0.43 *	0.18			
CaCO <sub>3</sub>	−0.15	0.10	0.09	−0.08	−0.14	0.29 *	−0.01	0.80 *		
Clay	0.07	0.09	−0.02	0.08	0.04	0.08	0.02	0.24	0.08	
GEC	−0.31 *	0.13	0.10	−0.03	−0.06	0.27 *	0.22	0.83 *	0.72 *	0.15

The highest correlation ( $r = 0.83$ ,  $p < 0.01$ ,  $n = 0.83$ ) was seen between GEC and SEC, followed by a significant correlation of 0.72 between GEC and soil CaCO<sub>3</sub> (Table 4). Among the topographical factors, elevation, LS factor showed a significant negative ( $r = -0.31$ ;  $p < 0.05$ ) and positive ( $r = 0.27$ ;  $p < 0.05$ ) correlation with groundwater EC. Besides, Flow length also had a slightly strong relationship with groundwater EC; however, the relationship was not significant. Furthermore, other topographical factors had no relationship with groundwater EC. Interestingly, clay particles showed a positive correlation with groundwater EC ( $r = 0.15$ ); however, it is not significant. Table 5 presents the MLR model's information in predicting the GEC by environmental data (soil-topographical parameters). The MLR showed that SEC, LS factor, CaCO<sub>3</sub>, Flow length, TWI, and clay and elevation were noticeably related to groundwater EC (Table 5).

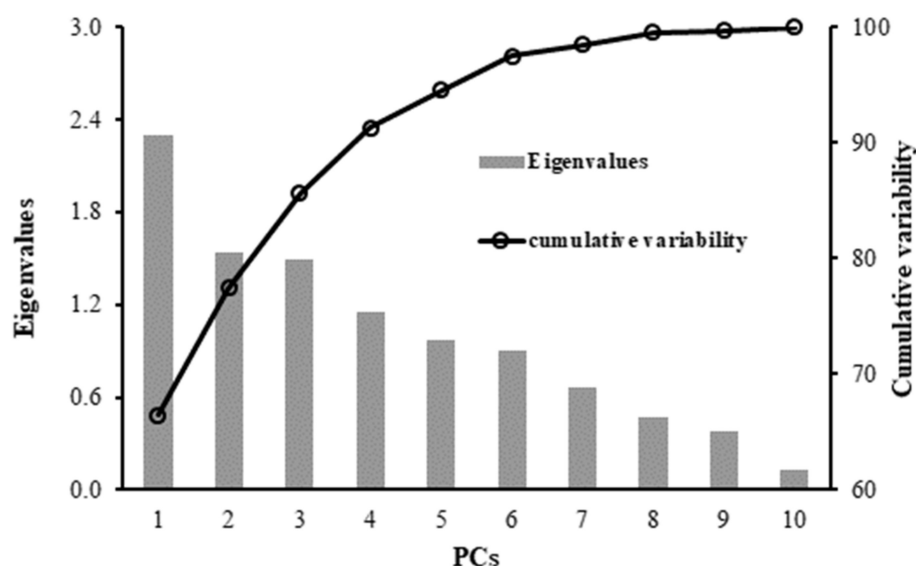
**Table 5.** Regression information for modeling groundwater EC.

Parameter	Standardized T Coefficients Beta	Std. Err. of B	Unstandardized Coefficients B	Significance Level ( <i>p</i> )	VIF
Intercept			27,248.24	0.030	3.5
Soil EC	0.78	0.08	0.56	0.000	2.6
LS factor	−0.10	0.05	−767.96	0.045	4.5
%CaCO <sub>3</sub>	0.21	0.07	97.04	0.005	3.4
Elevation	−0.09	0.04	−18.62	0.024	3.9
Flow length	0.095	0.04	2987.94	0.029	4.1
TWI	0.06	0.04	165.88	0.137	7.5
Clay	−0.05	0.04	−26.86	0.238	8.2

The LS factor's adverse effects were significant at  $p < 0.05$  and elevation at  $p < 0.05$ . The SEC's positive effects were significant at  $p < 0.0001$ , %CaCO<sub>3</sub> at  $p < 0.001$ , and Flow length at  $p < 0.05$ . The linear model is given as follows:

$$\text{GEC} = 27,248 + 0.56 \text{ SEC} - 768 \text{ LS factor} + 97 \text{ CaCO}_3 - 8 \text{ Elevation} + 1331 \text{ Flow length} \quad (11)$$

where GEC and SEC are groundwater and soil EC ( $\mu\text{Scm}^{-1}$ ), respectively, CaCO<sub>3</sub> is soil calcium carbonate (%), Elevation is altitude (m), and Flow length is the length of the flow (m). Figure 6 indicates the raw variance of data (Eigenvalues) for PCs. The first four PCs describe variance >90% (Figure 6). Coefficients of GEC with the first-three PCs indicate that PC1 was correlated with GEC ( $r = 0.80$ ;  $p < 0.01$ ), followed by PC2 with  $r = 0.38$ . PC2 was considered for 10.5% of the variance; it was specified by aspect, elevation, and LS factor (Table 6; Figure 6). PC3 with 8.2% of the whole variance was specified by TWI ( $r = 0.688$ ;  $p < 0.05$ ) and slope ( $r = 0.687$ ;  $p < 0.05$ ) (Figure 7b). Finally, PC4 was mainly characterized by flow direction ( $r = 775$ ;  $p < 0.05$ ) and clay ( $r = 535$ ;  $p < 0.05$ ) (Figure 7c).



**Figure 6.** Eigenvalues and cumulative variability are associated with the principal axes of the data matrix.

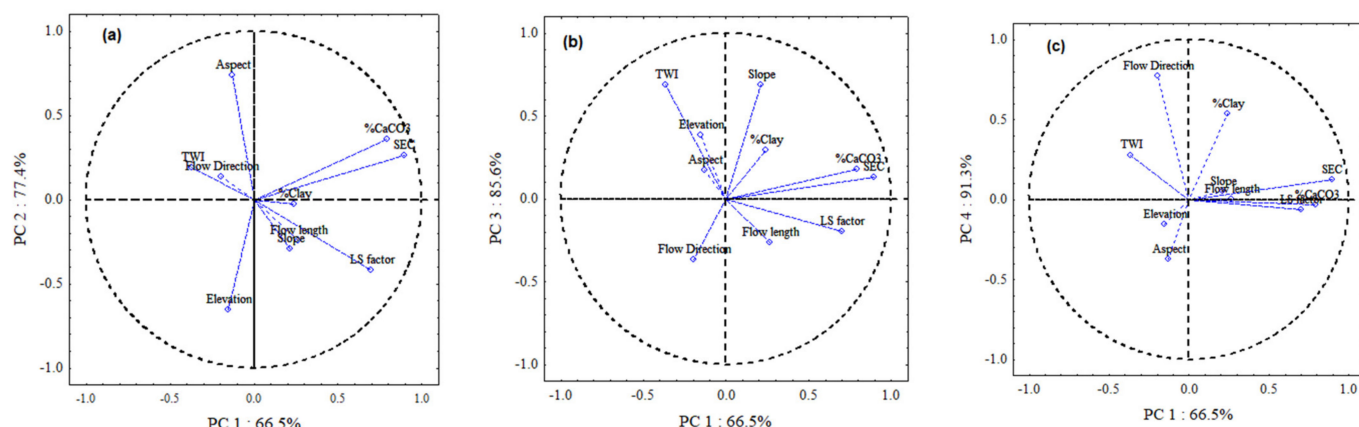


Figure 7. Correlation circles: PC1–PC2 (a), PC1–PC3 (b) and PC1–PC4 (c).

Table 6. Eigenvectors of the correlation matrix used to generate the new principal component (PC) variables.

	PC 1	PC 2	PC 3	PC 4
Elevation	−0.155	−0.651	0.382	−0.155
Slope	0.211	−0.291	0.687	0.054
Aspect	−0.128	0.739	0.174	−0.371
TWI	−0.368	0.191	0.688	0.279
Flow direction	−0.198	0.133	−0.365	0.775
LS factor	0.698	−0.421	−0.200	−0.062
Flow length	0.262	−0.243	−0.263	0.004
SEC	0.897	0.260	0.128	0.125
CaCO <sub>3</sub>	0.791	0.357	0.177	−0.033
Clay	0.240	−0.027	0.293	0.535

$R^2 = 0.92$  indicates that the developed model (Equation (11)) explains 92% of the variance. It shows that the regression predictions perfectly fit the data. Table 5 shows that soil EC with Beta = 0.78 is the leading property used to predict the groundwater EC values, followed by CaCO<sub>3</sub> and LS factor. Table 4 shows a very positive correlation ( $r = 0.83$ ) between groundwater EC and soil EC.

### 3.3. PCs Analysis

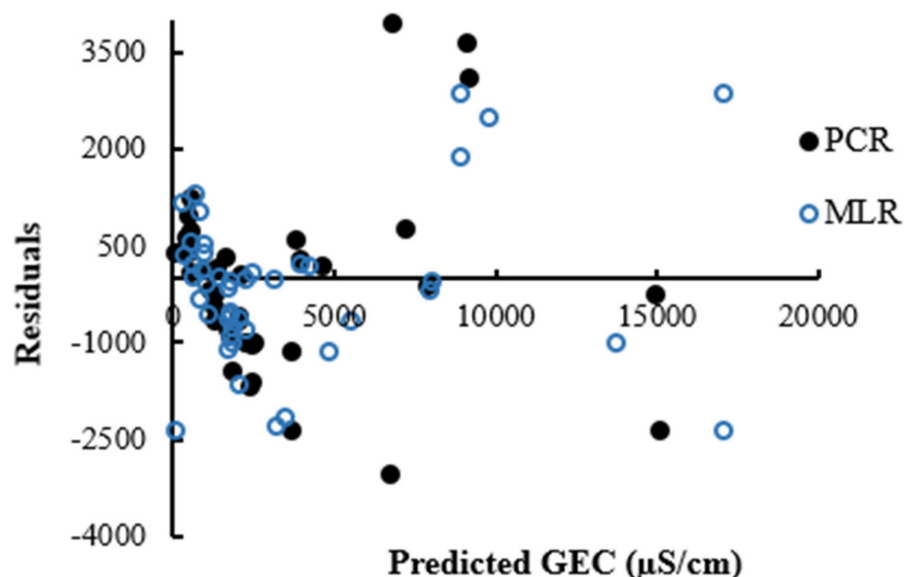
Figure 6 indicates the raw variance of data (Eigenvalues) for PCs. The first four PCs describe variance >90% (Figure 6). The PC1, indicating 66.5% of the variance, interprets the groundwater EC (Figure 7a). PC1 was characterized by soil EC ( $r = 0.897$ ;  $p < 0.01$ ) and CaCO<sub>3</sub> ( $r = 0.791$ ;  $p < 0.01$ ) and LS factor ( $r = 0.698$ ;  $p < 0.05$ ), which is consistent with the results of the MLR (Table 4). As mentioned above and given in Table 3, soil EC with Beta = 0.78, CaCO<sub>3</sub> with Beta = 0.21, and LS factor with Beta = 0.10 were the most important/effective parameters in the GEC prediction (Table 5). As soils in the study site are calcareous due to carbonate formation, CaCO<sub>3</sub> has a strong impact on groundwater quality [3,28].

Coefficients of GEC for the first-three PCs indicates that PC1 was correlated with GEC ( $r = 0.80$ ;  $p < 0.01$ ), followed by PC2 with  $r = 0.38$ . PC2 considered for 10.5% of the whole variance; it was specified by aspect, elevation and LS factor (Table 6; Figure 7b). As shown in Figure 7b, these factors were shorter than others, which indicates the importance and impacts of these factors on GEC. Interestingly, in Figure 7c, the mentioned factors, including aspect, elevation and LS factor, are represented as the most influential factors affecting the GEC. PC3 with 8.2% of the whole variance was specified by TWI ( $r = 0.688$ ;  $p < 0.05$ ) and slope ( $r = 0.687$ ;  $p < 0.05$ ) (Figure 7b). Finally, PC4 was mainly characterized by flow direction ( $r = 0.775$ ;  $p < 0.05$ ) and %clay ( $r = 0.535$ ;  $p < 0.05$ ) (Figure 7c).

The regression model to predict groundwater EC that used the four first PCs is presented as follows:

$$\text{GEC} = 3110.5 + 3021.6 \text{ PC1} + 1264.7 \text{ PC2} + 310.5 \text{ PC3} \quad R^2 = 0.86 \quad (12)$$

where GEC is groundwater EC. PC1 with Beta = 0.83 was the most important variable in the Equation (10), followed by PC2 with Beta = 0.32. Interestingly, although the first four PCs were considered the input variables, the PC4 did not appear in the model. Residual plot (Figure 8) illustrates that MLR and PCR models are well scattered around the horizontal axis, indicating appropriate fitted models. Positive values for the residual (on the y-axis) mean the prediction was too low, and negative values mean the prediction was too high; 0 means the guess was exactly correct. As shown in Figure 8, most of the points are around the 0 lines, indicating that the predictions were fine and acceptable. Figure 9 shows the predicted groundwater EC by MLR and PCR models versus observed groundwater EC in both development and validation data sets. In both data sets, the PCR model underestimated the groundwater EC by the largest amount (Figure 9a,b). However, a better agreement between observed and predicted GEC was achieved with the MLR (Figure 9a,b).



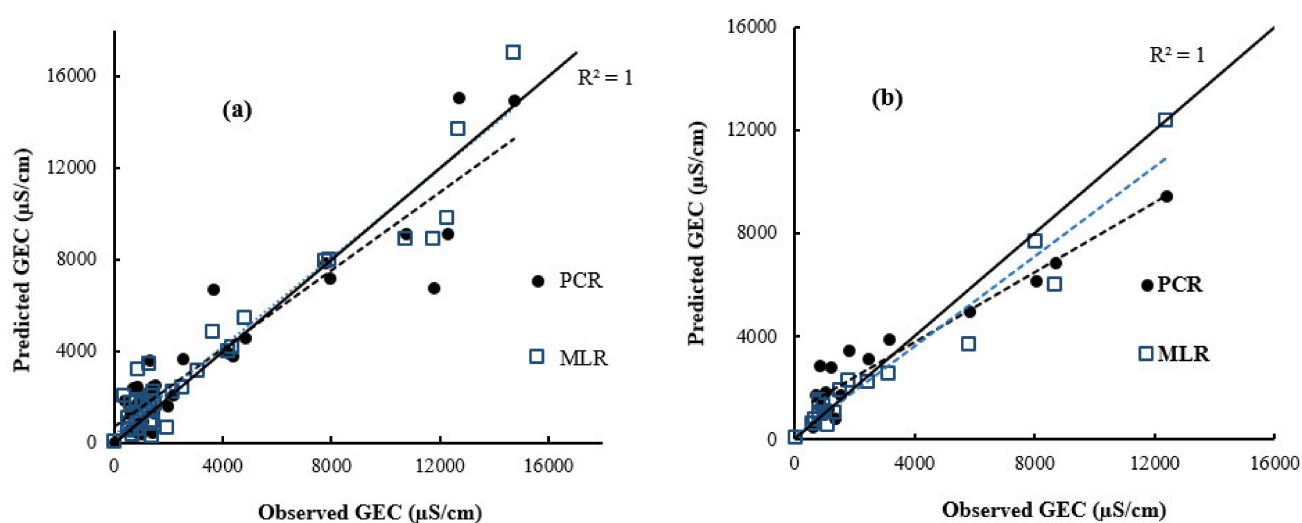
**Figure 8.** The residual plot in the prediction of groundwater EC (GEC) using multiple linear regression (MLR) and principal component regression (PCR).

The statistic criteria indices for estimating groundwater EC for MLR and PCR are given in Table 7. Figure 10 indicates no significant difference between MLR and PCR in both development and validation data sets.

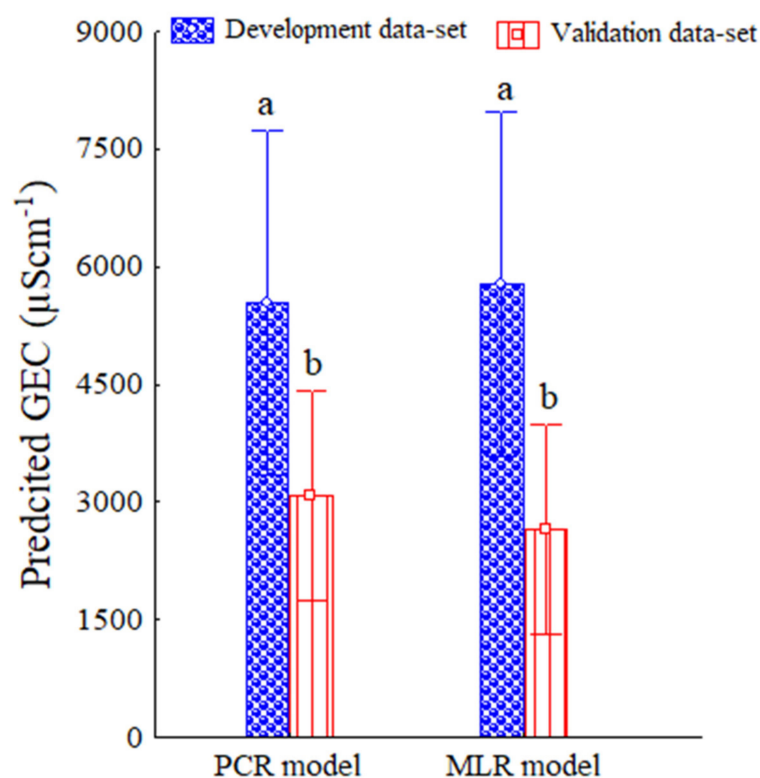
**Table 7.** Statistic criteria indices for estimating groundwater EC using MLR and PCR.

Methods	R <sup>2</sup>		RMSE (μScm <sup>−1</sup> )		ME (μScm <sup>−1</sup> )	
	Validation Data Set	Development Data Set	Validation Data Set	Development Data Set	Validation Data Set	Development Data Set
MLR	0.94	0.92	450	478	125	132
PCR	0.91	0.86	486	520	210	187





**Figure 9.** Observed groundwater EC (GEC) versus the predicted groundwater EC (GEC) for MLR and PCR in development (a) and validation (b).



**Figure 10.** Mean predicted groundwater EC by the MLR and PCR models in the validation and development data sets.

A spatial map of the GEC (shown in Figure 11) was created in GIS based on the developed MLR model (Equation (9)). The groundwater's EC value varied from  $43 \mu\text{Scm}^{-1}$  in the north to  $15,125 \mu\text{Scm}^{-1}$  in the south of the site (Figure 9). Figure 9 shows that only 16% of the Marvdasht Aquifer ( $317.5 \text{ Km}^2$ ) had EC higher than  $10,000 \mu\text{Scm}^{-1}$ , which is unsuitable for agriculture as the most critical activity Marvdasht Plain. However, more than 50% of the aquifer area had EC less than  $1000 \mu\text{Scm}^{-1}$ , suitable for agriculture (Figure 11).

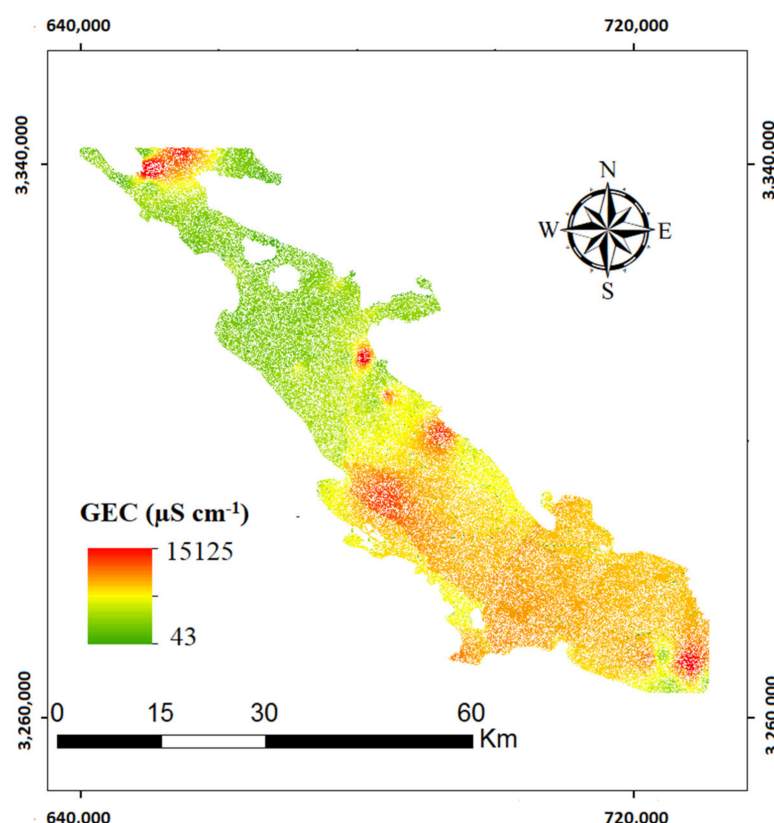


Figure 11. Spatial map of the groundwater EC (GEC) obtained from the MLR model.

#### 4. Discussion

The soils provided a wide range of soil particles, clay from 15.1 to 40.0%, sand from 30.1 to 55.0%, and silt from 24.0 to 40.0%. Clay with a CV value of 17.1 had the highest variability, followed by silt with a CV value of 10.1. The main soil texture in the study site was silty loam, which can transfer water easily to groundwater. Soils with a high amount of silt, which is the most sensitive to wind and water erosion [13], have weak and small aggregate, sequentially increasing the potential of soil erosion, consequently increasing fine particles, which can transport salt into the groundwater. A significant positive correlation was also found between organic matter (OM) with  $\text{CaCO}_3$  ( $r = 0.49$ ,  $p < 0.05$ ,  $n = 120$ ). Ostovari et al. [13] reported a correlation coefficient of 0.36 between OM with  $\text{CaCO}_3$  (not shown).  $\text{CaCO}_3$  has a high amount of  $\text{Ca}^{2+}$  that plays an important role in creating big and stable aggregates by acting as a binding agent for flocculating soil minerals, increasing soil aggregates resistance against runoff and raindrop detachment, which encourages the solvent materials to move down into the groundwater. Ostovari et al. [27,28] showed that  $\text{CaCO}_3$  had a significant influence on soil EC, indirectly affecting groundwater EC. This could be associated with carbonate formation with the high amount of soluble materials, which increase the salinity of groundwater. The type of water in the Marvdasht Aquifer is Na-Cl, which could be supported by a mean total dissolve solid (TDS) value of 2400.7 mg/L and EC of 4001.2  $\mu\text{S}/\text{cm}$ . Honarbakhsh et al. [3] also reported that the type of groundwater in Marvdasht is Na-CL. The  $\text{Mg}^{2+}$  is an important cation in this aquifer due to carbonate and marl formation that consists of large quantities of dolomite, which is in agreement with the findings of Honarbakhsh et al. [3]. Charging the Marvdasht Aquifer with the Kor River, which passes through the chalky-saline formation of Gachsaran, transports large amounts of urban sewage with high concentrations of  $\text{Cl}^-$ . Dissolution of carbonate formations is a simple and common weathering reaction in aquifers in semi-arid regions.

In addition, the positive correlation between clay and groundwater EC indicates that carbonate elements ( $\text{Ca}^{2+}$  and  $\text{Mg}^{+2}$ ) were absorbed on the clay surface, increasing ground-

water EC. The  $C_0/\text{sill}$  (i.e., 0.292–0.604) indicates a moderate dependency (Table 2) [32]. Besides, Table 3 gives the OK results to map soil parameters. It was highlighted that the OK provided good results in mapping soil items [30,32]. For groundwater EC, the spherical model was the best; a spatial dependency of 0.345 indicates a moderately spatial dependency for GEC. As shown in Figure 4c,d, GEC and SEC's spatial distribution is very similar, which agrees with the correlation between these parameters in this study site. The reason could be due to the existence of the soluble carbonate formation in the study site that leads to increased soil  $\text{CaCO}_3$  and groundwater EC, showing a significant correlation between SEC and soil  $\text{CaCO}_3$  ( $r = 0.81$ ,  $p < 0.01$ ,  $n = 75$ ), which is in line with finding by Ostovari et al. [1]. In addition, in the semi-arid regions, groundwater salinity may also attributed to the formation of salt layers by leaching from the soil surface due to high evaporation during the dry seasons.

Among topographical parameters, LS factor and elevation had a significant correlation of 27 and  $-31$  with groundwater EC, respectively (Table 3), which shows that increasing the elevation groundwater decreased EC. This could occur because precipitation increased with increasing elevation. As a result, the concentration of the soluble materials decreased. A positive correlation between LS and groundwater EC could indicate that groundwater EC increased with the increasing slope degree and slope length, due to high runoff from the higher slope downward, resulting in an increased salinity to groundwater because of water infiltration. In addition, the LS factor was correlated with aspect ( $r = -0.31$ ,  $p < 0.01$ ,  $n = 75$ ) and TWI ( $r = -0.44$ ,  $p < 0.01$ ,  $n = 75$ ). According to Equations (1) and (2), flow accumulation is an important parameter for calculating both TWI and LS factors. In addition, TWI significantly contributed to the slope ( $r = 0.26$ ,  $p < 0.01$ ,  $n = 75$ ), which indicates that the TWI index characterizes the impact of slope on the hydrological processes [28]. Naghibi et al. [18] and Haghizade et al. [15] reported a vital role of the LS factor and TWI in groundwater quality modeling. However, TWI and clay did not have a significance coefficient, and therefore, they did not appear in the developed MLR model. The  $\text{VIF} > 5$  for two variables (TWI and clay) indicated multicollinearity among input variables (Table 5). As aforementioned, the importance of TWI for the soil and water hydrological process has been widely reported [15,18] as it reflects the effects of several topographical factors that could have significant impacts on groundwater quality.

Based on the indices, the MLR had a better performance than PCR in predicting the GEC by using soil-topographical parameters because it had a higher  $R^2$  (0.92 and 0.94), lower RMSE (478 and  $450 \mu\text{Scm}^{-1}$ ) and ME (132 and  $125 \mu\text{Scm}^{-1}$ ) in both development and validation data sets, which is similar to the finding by Belkhiri and Narany [7]. Saleem et al. [33] showed that the MLR method could predict groundwater EC at a 95% significance level. Kamakshaiah and Seshadri [34] stated that the correlation between groundwater quality parameters provides an appropriate tool for assessing water quality. Ling et al. [8] reported the multivariate statistical model for water quality assessment in China. Yadav et al. [16] reported the high performance of combined GIS-PCA to evaluate groundwater quality in India. Elizabeth et al. [35] and Tahmassebipoor et al. [36] also showed the PCA model's superior performance in predicting the groundwater quality. However, Van-Westen et al. [37] illustrated that one of the most critical limitations of bivariate statistical models is that they use conditional independence assumptions. By the way, these models are related to groundwater potential locations and influential factors exceptionally. Another substantial disadvantage of bivariate statistical models, as presented by Pradhan et al. [38], is that we could not calculate each factor's importance and output maps' role. Therefore, the mentioned drawbacks can affect each finding in different cases.

Lee et al. [39,40] showed the benefit of GIS and statistical models for landslide evaluation. The acceptable results of using GIS coupled with statistical models for groundwater quality assessment in Iran have been reported [41–43]. As aforementioned, in the south of the Marvdasht Plain, groundwaters are affected by the soluble chalky-carbonate formation, leading to increased soluble salt (total dissolved solid). According to Ostovari et al. [1], the north of the aquifer is influenced by Doroudzan Dam, which was constructed on the Kor

River and leads to decreased salt concentration (EC) in this part of the aquifer, affecting plant growth [44]. The groundwater EC's spatial variability from the MLR model (Figure 9) is very similar to the soil EC spatial distribution (Figure 4c). The study agrees well with the flow computational investigations of Ostovari et al. [1]. It was revealed that the flow direction of groundwater in the Marvdasht Aquifer basin is toward the northeast in the southeastern part and toward the south in the northeastern part. The flow direction changes west in the eastern part and northwest in the southern side. On the western side, flow is generally toward the coast. In the present study, most of the very high and high groundwater potential zones were associated with the central zone, which coincides with the groundwater flow convergence area identified by Honarbakhsh et al. [3]. The Kappa index of 0.81 indicates a good match between the GEC map obtained by the MLR model and the GEC map obtained from the kriging method (Figure 4d). However, the GEC map predicted by the MLR model showed a little overestimation compared to the GEC map created by the EC data and interpolating method (Figure 4d). It is worth mentioning that there are some uncertainties in applying the statistical method for groundwater quality assessment. Uncertainty is the situation in which there is a lack of confidence about an event's specific outcome. Reasons for this lack of confidence might include a judgment of the information as incomplete, blurred, inaccurate, unreliable, inconclusive, or potentially false.

## 5. Conclusions

Groundwater is an invaluable water source in semi-arid regions like Iran. Therefore, the present work attempted to assess the Marvdasht groundwater quality potential by applying statistical models in GIS. The results showed that the groundwater EC had a strong significant correlation ( $r = 0.73$ ) with soil  $\text{CaCO}_3$ , indicating the effects of geological formation on groundwater quality parameters. The results showed that the highest relationship ( $r = 0.83$ ;  $p < 0.01$ ;  $n = 75$ ) was observed between groundwater EC and soil EC. According to the MLR result, soil EC, LS factor,  $\text{CaCO}_3$ , Flow length, TWI, and clay and elevation were significantly correlated to groundwater EC. PCR results showed that the first four PCs explained 91% of the variance, and soil EC characterized PC1 with  $r = 0.89$ ,  $\text{CaCO}_3$  with  $r = 0.79$ , and LS factor with  $r = 0.69$ . The MLR model with  $R^2 = 0.94$ ,  $\text{RMSE} = 450 \mu\text{Scm}^{-1}$  and  $\text{ME} = 125 \mu\text{Scm}^{-1}$  in the validation data set had a better prediction than the PCR model in modeling the groundwater EC. The spherical semi-variogram model showed high performance for all soil and topography attributes. The map obtained from the ordinary kriging method coupled with the MLR model showed that 16% of the aquifer is unsuitable for agriculture, located in the southern part of the study area predominated by soluble geological formations with a high amount of carbonate and salts. We suggest that it is better to use this part of the study area for other activities due to low groundwater quality, which does not need a high amount of groundwater. We conclude that a GIS-based statistical model and environmental data can be used for monitoring the groundwater quality potential in semi-arid regions of Iran. We recommend using quality data of more wells and using other statistical methods for modeling the groundwater quality. We suggest testing the methodology's validity in this paper with a bigger data set in the different watersheds with different soil and topography conditions. We recommend using new data-mining techniques such as random forest to determine the most influential soil parameters and topography attributes and develop a simple linear model according to the selected important parameters.

**Author Contributions:** Formal analysis, A.A.; Investigation, A.A.; Supervision, A.A.; Writing—original draft, A.A. and N.K. and R.K.; Methodology, E.N.; Resources, E.N., and A.S.; Data curation, E.N.; Software, A.S.; Writing—review & editing, A.S. and S.M.S.A.; Validation, N.K.; Project administration, N.K.; Visualization, A.D. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not Applicable.



**Informed Consent Statement:** Not Applicable.

**Data Availability Statement:** Not Applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Ostovari, Y.; Beigi-Harchegani, H.; Asgari, K. A fuzzy logic approach for assessment and mapping of groundwater irrigation quality: A case study of Marvdasht aquifer, Iran. *Arch. Agron. Soil Sci.* **2014**, *61*, 711–723. [\[CrossRef\]](#)
- Haghizadeh, A.; Moghaddam, D.D.; Pourghasemi, H.R. GIS-based bivariate statistical techniques for groundwater potential analysis (an example of Iran). *J. Earth Syst. Sci.* **2017**, *126*, 109. [\[CrossRef\]](#)
- Honarbaksh, A.; Azma, A.; Nikseresht, F.; Mousazadeh, M.; Eftekhari, M.; Ostovari, Y. Hydro-chemical assessment and GIS-mapping of groundwater quality parameters in semi-arid regions. *J. Water Supply Res. Technol.* **2019**, *68*, 509–522. [\[CrossRef\]](#)
- Honarbaksh, A.; Tahmoures, M.; Tashayo, B.; Mousazadeh, M.; Ingram, B.; Ostovari, Y. GIS-based assessment of groundwater quality for drinking purpose in northern part of Fars province, Marvdasht. *J. Water Supply Res. Technol.* **2019**, *68*, 187–196. [\[CrossRef\]](#)
- Chen, W.; Li, H.; Hou, E.; Wang, S.; Wang, G.; Panahi, M.; Li, T.; Peng, T.; Guo, C.; Niu, C.; et al. GIS-based groundwater potential analysis using novel ensemble weights-of-evidence with logistic regression and functional tree models. *Sci. Total Environ.* **2018**, *634*, 853–867. [\[CrossRef\]](#)
- Lee, S.; Hyun, Y.; Lee, M.-J. Groundwater Potential Mapping Using Data Mining Models of Big Data Analysis in Goyang-si, South Korea. *Sustainability* **2019**, *11*, 1678. [\[CrossRef\]](#)
- Belkhiri, L.; Narany, T.S. Using multivariate statistical analysis, geostatistical techniques and structural equation modeling to identify spatial variability of groundwater quality. *Water Res. Manag.* **2015**, *29*, 2073–2089. [\[CrossRef\]](#)
- Ling, T.-Y.; Soo, C.-L.; Liew, J.-J.; Nyanti, L.; Sim, S.-F.; Grinang, J. Application of Multivariate Statistical Analysis in Evaluation of Surface River Water Quality of a Tropical River. *J. Chem.* **2017**, *2017*, 5737452. [\[CrossRef\]](#)
- Zhang, Y.; Xu, M.; Li, X.; Qi, J.; Zhang, Q.; Guo, J.; Yu, L.; Zhao, R. Hydrochemical Characteristics and Multivariate Statistical Analysis of Natural Water System: A Case Study in Kangding County, Southwestern China. *Water* **2018**, *10*, 80. [\[CrossRef\]](#)
- Ho, R. *Handbook of Univariate and Multivariate Data Analysis and Interpretation with SPSS*; CRC Press: Boca Raton, FL, USA, 2006.
- Wieczorek, W.F.; Delmerico, A.M. Geographic information systems. *Wiley Interdiscip. Rev. Comput. Stat.* **2009**, *1*, 167–186. [\[CrossRef\]](#)
- Shahid, S.; Nath, S.K.; Kamal, A.S.M. GIS Integration of Remote Sensing and Topographic Data Using Fuzzy Logic for Ground Water Assessment in Midnapur District, India. *Geocarto Int.* **2002**, *17*, 69–74. [\[CrossRef\]](#)
- Ostovari, Y.; Moosavi, A.A.; Pourghasemi, H.R. Soil loss tolerance in calcareous soils of a semiarid region: Evaluation, prediction, and influential parameters. *Land Degrad. Dev.* **2020**, *31*, 2156–2167. [\[CrossRef\]](#)
- Rahmati, O.; Samani, A.N.; Mahdavi, M.; Pourghasemi, H.R.; Zeinivand, H. Groundwater potential mapping at Kurdistan region of Iran using analytic hierarchy process and GIS. *Arab. J. Geosci.* **2015**, *8*, 7059–7071. [\[CrossRef\]](#)
- Arulbalaji, P.; Padmalal, D.; Sreelash, K. GIS and AHP Techniques Based Delineation of Groundwater Potential Zones: A case study from Southern Western Ghats, India. *Sci. Rep.* **2019**, *9*, 1–17. [\[CrossRef\]](#)
- Yadav, K.K.; Gupta, N.; Kumar, V.; Choudhary, P.; Khan, S.A. GIS-based evaluation of groundwater geochemistry and statistical determination of the fate of contaminants in shallow aquifers from different functional areas of Agra city, India: Levels and spatial distributions. *RSC Adv.* **2018**, *8*, 15876–15889. [\[CrossRef\]](#)
- Naghibi, S.A.; Pourghasemi, H.R.; Dixon, B. GIS-based groundwater potential mapping using boosted regression tree, classification and regression tree, and random forest machine learning models in Iran. *Environ. Monit. Assess.* **2016**, *188*, 1–27. [\[CrossRef\]](#) [\[PubMed\]](#)
- Mosaferi, M.; Pourakbar, M.; Shakerkhatibi, M.; Fatehifar, E.; Belvasi, M. Quality modeling of drinking groundwater using GIS in rural communities, northwest of Iran. *J. Environ. Heal. Sci. Eng.* **2014**, *12*, 99. [\[CrossRef\]](#) [\[PubMed\]](#)
- Abdalla, F.; Moubark, K.; Abdelkareem, M. Groundwater potential mapping using GIS, linear weighted combination techniques and geochemical processes identification, west of the Qena area, Upper Egypt. *J. Taibah Univ. Sci.* **2020**, *14*, 1350–1362. [\[CrossRef\]](#)
- Ijumulana, J.; Ligate, F.; Bhattacharya, P.; Mtalo, F.; Zhang, C. Spatial analysis and GIS mapping of regional hotspots and potential health risk of fluoride concentrations in groundwater of northern Tanzania. *Sci. Total Environ.* **2020**, *735*, 139584. [\[CrossRef\]](#)
- Sangab Zagros. *Report of the Fars Regional Water Authority. Detailed Project on the Study of Groundwater of Marvdasht-Kharame Plain*; Sangab Zagros: Shiraz, Iran, 2009; p. 225.
- USDA. *Keys to Soil Taxonomy*, 11th ed.; USDA National Resources Conservation Service: Washington, DC, USA, 2010.
- Charman, P.E.V.; Murphy, B.W. Soils (their properties and management). In *Land and Water Conservation*, 2nd ed.; Oxford University Press: South Melbourne, Australia, 2000; pp. 206–212.
- Ostovari, Y.; Ghorbani-Dashtaki, S.; Bahrami, H.-A.; Naderi, M.; Dematte, J.A.M.; Kerry, R. Modification of the USLE K factor for soil erodibility assessment on calcareous soils in Iran. *Geomorphology* **2016**, *273*, 385–395. [\[CrossRef\]](#)
- Webster, R.; Oliver, M.A. *Geostatistics for Environmental Scientists*; Wiley: Hoboken, NJ, USA, 2007.
- Moore, I.D.; Burch, G.J. Physical Basis of the Length-slope Factor in the Universal Soil Loss Equation. *Soil Sci. Soc. Am. J.* **1986**, *50*, 1294–1298. [\[CrossRef\]](#)



27. Ostovari, Y.; Ghorbani, S.; Bahrami, H.; Naderi, M.; Dematte, J.A.M. Soil loss prediction by an integrated system using RUSLE, GIS and remote sensing in semi-arid region. *Geoderma Region.* **2017**, *11*, 28–36. [[CrossRef](#)]
28. Ostovari, Y.; Ghorbani-Dashtaki, S.; Kumar, L.; Shabani, F. Soil erodibility and its prediction in semi-arid regions. *Arch. Agron. Soil Sci.* **2019**, *65*, 1688–1703. [[CrossRef](#)]
29. Mirzaee, S.; Ghorbani-Dashtaki, S.; Mohammadi, J.; Asadi, H.; Asadzadeh, F. Spatial variability of soil organic matter using remote sensing data. *Catena* **2016**, *145*, 118–127. [[CrossRef](#)]
30. Hassoun, M.H. *Fundamentals of Artificial Neural Networks*; MIT Press: Cambridge, MA, USA, 1995.
31. Mirzaee, S.; Ghorbani-Dashtaki, S.; Kerry, R. Comparison of a spatial, spatial and hybrid methods for predicting inter-rill and rill soil sensitivity to erosion at the field scale. *Catena* **2020**, *188*, 104439. [[CrossRef](#)]
32. Nikseresht, F.; Honarbakhsh, A.; Ostovari, Y.; Afzali, S.F. Model Development to Predict CEC Using the Intelligence Data Mining Approaches. *Commun. Soil Sci. Plant Anal.* **2019**, *50*, 2178–2189. [[CrossRef](#)]
33. Saleem, A.D.; Igi, M.N.; Kumar, K.V. Correlation-regression model for physico-chemical quality of groundwater in the South Indian city of Gulbarga. *Afr. J. Environ. Sci. Technol.* **2012**, *6*, 353–364. [[CrossRef](#)]
34. Kamakshaiah, K.; Seshadr, R. Analysis of groundwater quality using regression model. *Inter. J. Engine Res.* **2015**, *3*, 6137–6149.
35. Celestino, A.E.M.; Cruz, D.A.M.; Sánchez, E.M.O.; Reyes, F.G.; Soto, D.V. Groundwater Quality Assessment: An Improved Approach to K-Means Clustering, Principal Component Analysis and Spatial Analysis: A Case Study. *Water* **2018**, *10*, 437. [[CrossRef](#)]
36. Tahmassebpour, N.; Rahmati, O.; Noormohamadi, F.; Lee, S. Spatial analysis of groundwater potential using weights-of-evidence and evidential belief function models and remote sensing. *Arab. J. Geosci.* **2016**, *9*, 1–18. [[CrossRef](#)]
37. Van Westen, C.J.; Rengers, N.; Terlien, M.T.J.; Soeters, R. Prediction of the occurrence of slope instability phenomena through GIS-based hazard zonation. *Geol. Rundsch.* **1997**, *86*, 404–414. [[CrossRef](#)]
38. Pradhan, A.M.S.; Dawadi, A.; Kim, Y.T. 2013 Use of different bivariate statistical landslide susceptibility methods: A case study of Kulekhani watershed Nepal. *J. Nepal Geol. Soc.* **2013**, *44*, 1–12.
39. Lee, S.; Lee, C.-W. Application of Decision-Tree Model to Groundwater Productivity-Potential Mapping. *Sustain. J. Rec.* **2015**, *7*, 13416–13432. [[CrossRef](#)]
40. Lee, S.; Hong, S.-M.; Jung, H.-S. A Support Vector Machine for Landslide Susceptibility Mapping in Gangwon Province, Korea. *Sustain. J. Rec.* **2017**, *9*, 48. [[CrossRef](#)]
41. Barzegar, R.; Moghaddam, A.A.; Deo, R.; Fijani, E.; Tziritis, E. Mapping groundwater contamination risk of multiple aquifers using multi-model ensemble of machine learning algorithms. *Sci. Total. Environ.* **2018**, *621*, 697–712. [[CrossRef](#)] [[PubMed](#)]
42. Golkarian, A.; Naghibi, S.A.; Kalantar, B.; Pradhan, B. Groundwater potential mapping using C5.0, random forest, and multivariate adaptive regression spline models in GIS. *Environ. Monit. Assess.* **2018**, *190*, 149. [[CrossRef](#)] [[PubMed](#)]
43. Pradhan, B.; Mousavi, S.M.; Golkarian, A.; Naghibi, S.A.; Kalantar, B. GIS-based Groundwater Spring Potential Mapping Using Data Mining Boosted Regression Tree and Probabilistic Frequency Ratio Models in Iran. *AIMS Geosci.* **2017**, *3*, 91–115. [[CrossRef](#)]
44. Safari, M.; Mousavi-Fard, S.; Nejad, A.R.; Sorkheh, K.; Sofo, A. Exogenous salicylic acid positively affects morpho-physiological and molecular responses of *Impatiens walleriana* plants grown under drought stress. *Int. J. Environ. Sci. Technol.* **2021**, 1–16. [[CrossRef](#)]